PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Audiovisual simulation inside the residential rooms of roadside buildings

Takumi ASAKURA[1]; Riku HASHIMOTO

[1] Tokyo University of Science, Japan

ABSTRACT

The effect of the auditory and visual stimuli on the subjective impression of the pass-by noise of running vehicle transmitted into the residential spaces was investigated based on the binaural sound auralization technieque and the three-dimensional visualization technique. As a basic study, the pass-by sound of the running vehicle transmitted into the rooms via the façade were simulated by using the geometrical acoustic simulation with relatively low computational costs. The visual condition inside the evaluated rooms were simulated by using three-dimensional modeling software, Unity. Then, the effect of the auditory and visual stimuli on the impression of the environments inside rooms were investigated by performing the subjective evaluation experiment based on the audiovisual simulation scheme. The experimental results are detailedly discussed.

Keywords: Auralization, Sound insulation, Transmission, Sound impression

1. INTRODUCTION

To evaluate the quality of the residential buildings, the sound environment is the important issue among various kinds of factors. The sound environments inside residential spaces can be kept to be silent by improving the sound insulation performance of houses. Many researches concerning the improvement of the sound insulation of buildings have been performed. The sound environments can be evaluated based on the quantitative ratings related to physical values, such as the sound reduction index of walls, however the subjective impression of the sound environment may not be judged by only the acoustic physical values but also other conditions of the spaces such as the visual conditions inside and outside rooms. There have been various kinds of researches which treated relationship between the auditory and visual perception of the environments [1-9]. Based on such like these knowledge, the subjective evaluation of sound environments should be performed by considering not only the acoustic issues but also visual ones in order to accurately predict the total human evaluation of environments. However, the sound insulation performance of residential buildings are basically evaluated by using the quantitative evaluation methods obtained from physical measurement or simulation, while the sound environments may also be effected by various kinds of other environmental factors including the visual stimuli. Especially, in recent researches, there has not been seen any obtainings related to the abovementioned issue of the effect of the visual factor on the subjective evaluation of sound transmitted into the residential spaces.

To investigate the effect of such sound environment on human, the sound data, which are recorded in in-situ environmental conditions, are reproduced in the laboratory such as the anechoic chamber. Such a recorded sound provide accurate simulation of the sound environment, while it requires much costs to acquire the recorded data. Then, the sound data can also be acquired by using auralization simulation techniques which are widely applied in the various kinds of fields related to the building acoustics[10-16]. As for the auralization of the traffic noise, the simulation technique of the vehicle running sound has been performed [15], and applied to simulation of road traffic noise. However, in this simulation, the sound insulation of the building façade are not considered which are important to evaluate the transmitted sound inside rooms. To cope with such a problem, we proposed the

[1] gf@aaa.com

auralization method [17] combining the measured vehicle sound and the sound-insulation characteristics obtained by wave-based vibroacoustic numerical method, which enabled to accurately simulate the vehicle sound transmitted to the rooms. However, the calculation of the transmission characteristics by the numerical scheme requires some computational costs.

In this paper, the effect of the auditory and visual stimuli on the subjective impression of the environment inside residential spaces was investigated as follows. Firstly, as a basic study, the vehicle running sound which transmits into the rooms via the façade are simulated by using the geometrical acoustic simulation with relatively low computational costs. In contrast, the visual condition inside the rooms were simulated by using three-dimensional modeling software. Then, the effect of the auditory and visual stimuli on the impression of the environments inside rooms were investigated by performing the subjective evaluation experiment based on the abovementioned simulation results.

2. SIMULATION METHOD OF PASS-BY SOUND OF VEHICLE

2.1 Procedure of auralization

The auralization scheme of the running pass-by sound of a vehicle is indicated as follows. Firstly, the dry source of the pass-by sound of running vehicle was recorded by using an omnidirectional microphone (the sound level meter, Rion, NL-32). The in-situ situation of the recorded place, and the spatial relationship between the running lane and the receiving points are shown in Fig. 1(a) and (b), respectively.

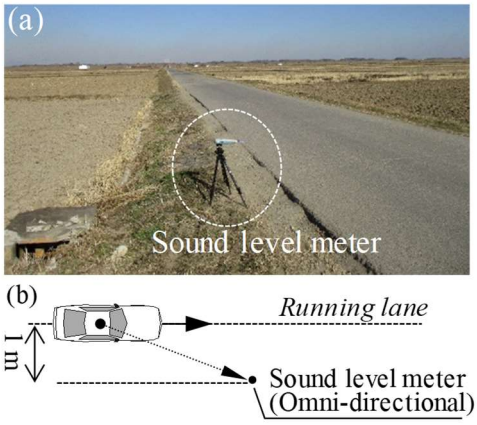

Figure 1 – Measurement site of the dry source.

At that place, there were not any buildings which reflect the pass-by sound. In order to record the pass-by sound with higher signal-to-noise ratio, the microphone was located at one-meter distance from the lane. Those data were obtained in the situation that a passenger car (TOYOTA, Mark II) was running in the road. The running velocity was maintained to be 40 km/h throughout the recording. The measured waveform is shown Fig. 2(a). Comparing to such a situation that the vehicle pass by near point to the receiving point, the sound is naturally more largely attenuated when the vehicle runs at further point from the receiving point. However, considering the usability of the dry source to auralization, the dry source should be indicating constant signal level excluding the distance attenuation. For that reason, the distance attenuation observed in the originally recorded dry source $p(t)$ was canceled, and modified into the steady-state dry source $p_{\mathrm{corr}}(t)$.

$$p_{\mathrm{corr}}(t) = p(t)\left|\frac{vt}{3600 \cdot d}\right| \tag{1}$$

Here v is the running velocity [km/h] of the vehicle (v = 40 in this study), d is the distance between the running lane and the receiving point (d = 0.001 km in this study), and t is the standardized time as the time of the vehicle passing by just in front of the receiving point is made to be 0 s. It should be noted that this correction is performed by assuming the running vehicle as a point source. The modified waveform $p_{\mathrm{corr}}(t)$ is indicated in Fig. 2(b). Another correction is also necessary because the pass-by sound also includes the Doppler effect which may affect the naturalness of the auralized sounds. For that reason, the modified waveform (Fig. 2(b)) and another waveform of Fig. 2(c), which was made by turning over the waveform of Fig. 2(b) in the time domain, were added. As a result, the waveform of Fig. 2(d) excluded the effect of the distance attenuation and the Doppler effect. Then, in order to generate the sound wave which propagates to the surface of the façade, the impulse responses, which were obtained by calculating the sound propagation from each point discretely arranged on the running lane (Fig. 3(a)) to the center point of the window (Fig.

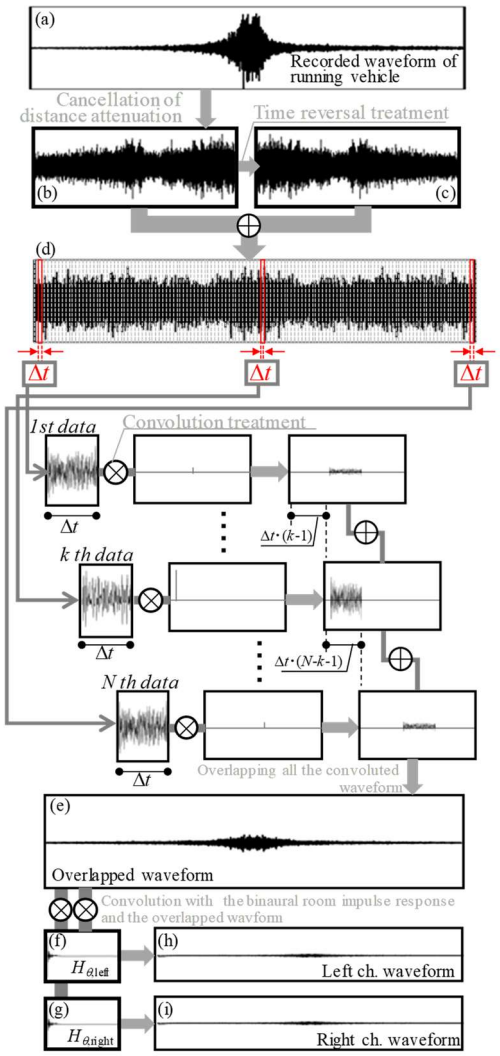

Figure 2 – Auralization scheme.

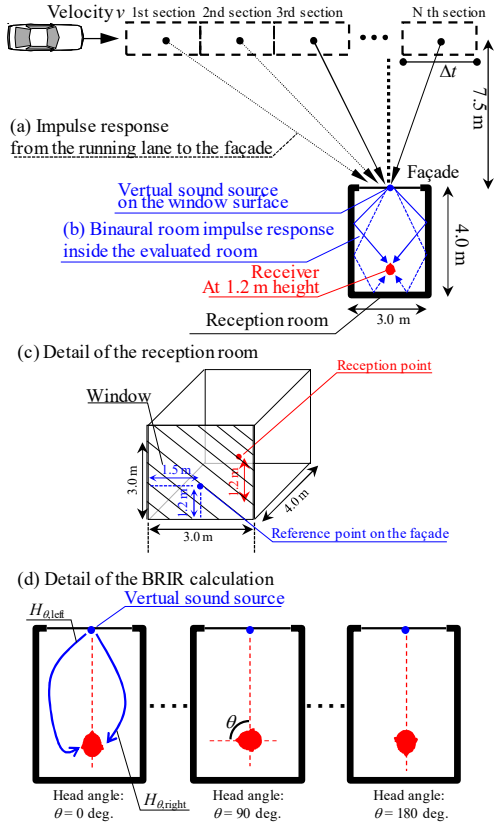

Figure 3 – Simulated conditions.

Table 1 – Simulated absorption coeffients.

Frequency [Hz]	125	250	500	1k	2k	4k
Surface A [-]	0.12	0.30	0.65	0.80	0.80	0.85
Surface B [-]	0.02	0.02	0.02	0.05	0.05	0.08

3(c)), were convoluted with each segment of the steady-state waveform as shown in the flowchart of Fig. 2(d). It should be noted that the waveform is divided into each segment with duration of Δt = 0.05 s in this study. The convoluted segments were overlapped to obtain resynthesized waveform of Fig. 2(e). Then, in order to consider the effect of the sound-insulation performance of façade, the waveform of Fig. 2(e) was attenuated by equalizing them with the frequency characteristics ΔL_2 of Fig. 4. Then, such a condition with the door open was also treated by simply using the unattenuated sound which is obtained by equalizing them with the flat frequency characteristics ΔL_1. The former waveform is hereafter called Wi_Op, whereas the latter one is ΔL_2 Wi_Cl. It should be noted the frequency characteristics ΔL_1 was set by referring the sound transmission loss of a window in which a glass pane of 3-mm-thick and 1.3 m×1.5 m is inserted in the aluminum sash.

Then, as a final step, the transmission characteristics from the window (the reference point on the window, Fig. 3(c)) to the receiving point inside the evaluated room, which include the room reverberation, were additionally considered to the attenuated waveform. In order to increase the presence of the reproduced sound field, binaural room impulse response (hereafter called BRIR) were calculated also in the CATT-acoustics. The detailed condition of the size of the room, the sound source, and receiving point are shown in Fig. 3(c). In addition, the sound absorption coefficients inside the room set in the simulation is shown as Surface A in

Table 1, which models such a nonreverberant room including carpets and other soft materials. However, by considering that the subjects see the omnidirectional view inside the room via the head-mounted display by turning their heads, the direction of the incoming sounds reproduced by the headphone needs to be unchanged by the head rotation of them. For that reason, the binaural reproduction system [18], which can dynamically and instantaneously change the BRIR linked to the head rotation, was adopted. This method has been validated in our previous paper for the auralization of the environmental sounds outside the rooms, however that has not been yet adopted for room-acoustic auralization. While the detailed method of the dynamic binaural reproduction system is the same as the reference [18], it requires 36 BRIRs in total, for each angle of the head at 10 deg. interval. Then, the 36 BRIRs were calculated on the CATT-acoustics by rotating the head direction as shown in Fig. 3(d). By using each of the BRIR, the stereo signal for the left and right channel as shown in Fig. 2(h) and (i) can be generated by convoluting the BRIR and the waveform of Fig. 2(e). As shown in Fig. 5, 36 binaural sound tracks are simultaneously loaded on the memory of the PC, and the reproduced sound tracks are continuously chosen based on the azimuth angle of the head monitored by the gyro sensor on the head mount device. By adopting the production system, the reproduced channel in accordance with momentarily changing direction of the head is automatically changed by using the 3-D integrated development environment software of Unity coded by the Visual C#.

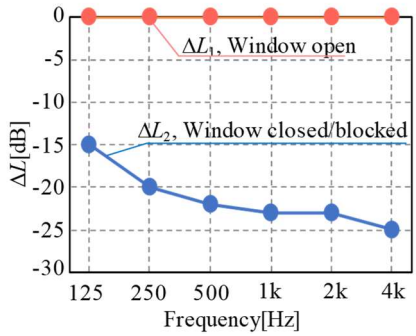

Figure 4 – Simulated attenuation characteristics of the façade.

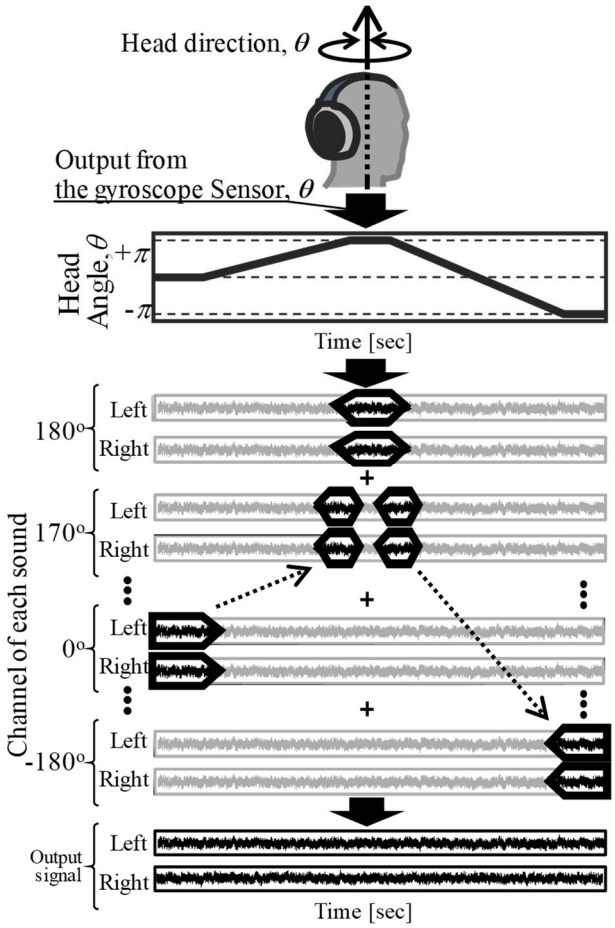

Figure 5 – Overview of the dynamic binaural simulation.

2.2 Procedure of visualization

Assuming an apartment for single persons built along a vehicle road, the visual environments inside the following three rooms (Fig. 6) were simulated; Type-1 is the room on the first floor, Type-2 is that on the third floors, and Type-3 is the room whose window is blocked. Then, the spatial relationship between the vehicle road and the apartment is shown in Fig. 6(a). These visual stimuli were simulated by using the Unity. These visual stimuli are reproduced by the head-mounted display, and presented to the subjects. Then, the visual stimuli is provided as omnidirectional visions which can be seen from the subject by rotating their head for left and right. In the visual condition of Type-3 (d), the opening of window is completely blocked, and the scenery of the exterior space cannot be seen, while those of (b) Type-1 and (c) Type-2 are not blocked. It should be noted that only the visual stimulus of (b) Type-1 provides moving image of the running vehicle, while the others provide still one because the running vehicle cannot be seen in the conditions of (c) and (d).

The audiovisual data, generated by combining these visual data with the auralized pass-by sound, were used as test stimuli in the following subjective evaluation experiment. Then, the simulated sound data for the vehicle running at the speed of 40 km/h was synchronously reproduced together with the visual stimuli. However, the synchronization can be confirmed as audiovisual stimuli only in the condition of Type-1, in which the running vehicle can be seen from the window. In addition, it is natural that the transmitted sound of the running vehicle into the rooms of on the first (Type-1) and third (Type-2) floor have different time and frequency characteristics due to the difference of the spatial relationship between the rooms and the sound source of the running vehicle. However, the main purpose of this study is to clarify the subjective influence of the visual stimuli on the sound impression. For that reason, the same sound data were used to both the audiovisual simulation of the condition Type-1, -2, and -3.

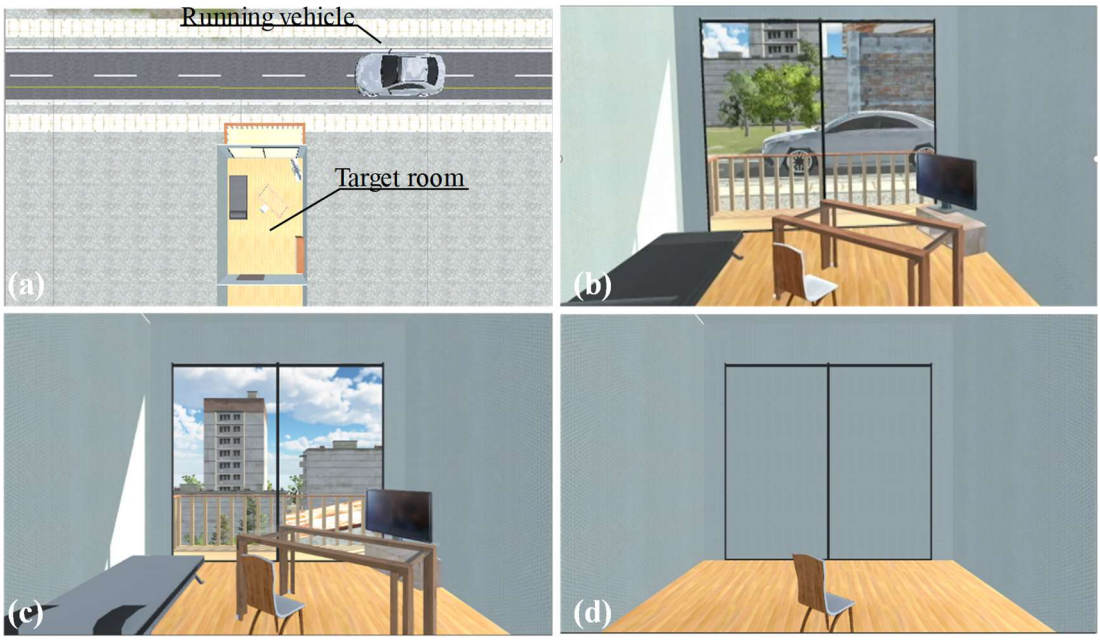

Figure 6 – Investigated visual conditions.

3. SUBJECTIVE EVALUATION EXPERIMENT

3.1 Scheme of subjective evaluation experiment

Firstly, the overall scheme of the experiment is described. In this experiment, the audiovisual stimuli inside rooms were simulated, and presented to the subjects. The experimental setting is shown in Fig. 7. As shown in the figure, to present these data to the subjects, a headphone (AUDIO-TECHINICA, ATH-W1000Z) and a head-mounted display (hereafter called HMD, OCULUS, Oculus Rift cv1) shown in Fig. 7 are used. Nine male subjects in the age of twenties participated in the experiment.

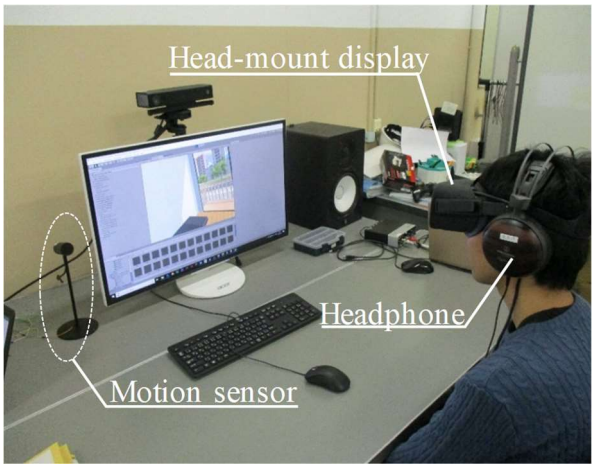

Figure 7 – Experimental settings of the subjective evaluation test.

The subjective evaluation experiment was performed to evaluate the impression of the transmitted sound of running vehicle into rooms by the semantic differential (SD) method. In this experiment, the 13 adjectives were adopted to evaluate the transmitted sound of running vehicle into rooms. To evaluate the extent of each adjectives, mono-polar seven categories from "Strongly disagree" to "Strongly agree" were adopted. Before the experiment, it was sufficiently instructed to each subject, to evaluate each condition independently, and not to evaluate each condition by comparing with each other. To familiarize the subjects with the SD method, randomly chosen conditions were additionally evaluated by the subjects as a trial. It should be noted that the results of the trial evaluation were excluded from the analysis and discussion. In the main experiment, all of the audiovisual stimulus conditions were randomly presented to the subjects. The subjects evaluated and wrote down their extent of impression of each adjective in the experimental sheet with sufficient length of time for writing. In the first experiment, the subjects were presented the six audiovisual data shown in Table 2, which are combination of the two audio data including the sounds filtered with the filter of ΔL_1 (Wi_Op) and ΔL_2 (Wi_Cl), and three visual data of Type-1, -2, and -3, respectively. As indicated in the table, each of the sound data has the equivalent continuous sound pressure level, L_{eq}, 65 dB for the condition of Wi_Op, and 53 dB for that of Wi_Cl. Each audiovisual data has time duration of 60 seconds, in which the pass-by sound of running vehicle is presented to the subject for five times, and the reproduction level of the sound was controlled that the equivalent continuous sound pressure level, L_{eq}, satisfies 65 dB and 53 dB for each condition of the sound.

Table 2 – Investigated conditions.

Condition	Sound stimulus	Visual stimulus	L_{Aeq} [dB]
So_V1	Wi_Op	Type-1	65.0 dB
Sc_V1	Wi_Cl	Type-1	53.0 dB
So_V2	Wi_Op	Type-2	65.0 dB
Sc_V2	Wi_Cl	Type-2	53.0 dB
So_V3	Wi_Op	Type-3	65.0 dB
Sc_V3	Wi_Cl	Type-3	53.0 dB

3.2 Results and discussion

The experimental results obtained by the SD method is shown as profile ratings in Fig. 8. These figures show the averaged values of all the subject's rating scores.

Firstly, Fig. 8(1) shows the evaluated results of So_V1, So_V2, and So_V3. The characterstic difference between each condition is supposed to indicate the effect of additionally presented scenery inside the evaluated room on the sound impression of the pass-by sound of the condition Wi_Op. As can be seen in the figure, the ratings vary depending on the visual stimuli. In the visual condition of Type-1, the impressions of "Powerful", "Agitated", and "Harsh" indicate lower ratings than the conditions of Type-2, whereas, in other adjectives, those of Type-1 indicate higher ratings than those of Type-2.

Secondly, Fig. 8(2) shows the evaluated results of So_V1, So_V2, and So_V3. The ratings of relatively

negative impression (“Powerful”, “Agitated”, and “Harsh”) indicate higher ratings, whereas those of positive impressions other than the abovementioned three adjectives indicate lower ratings. The changes of the ratings into more positive impression may be effected by the decrease of the sound pressure level of the pass-by noise.

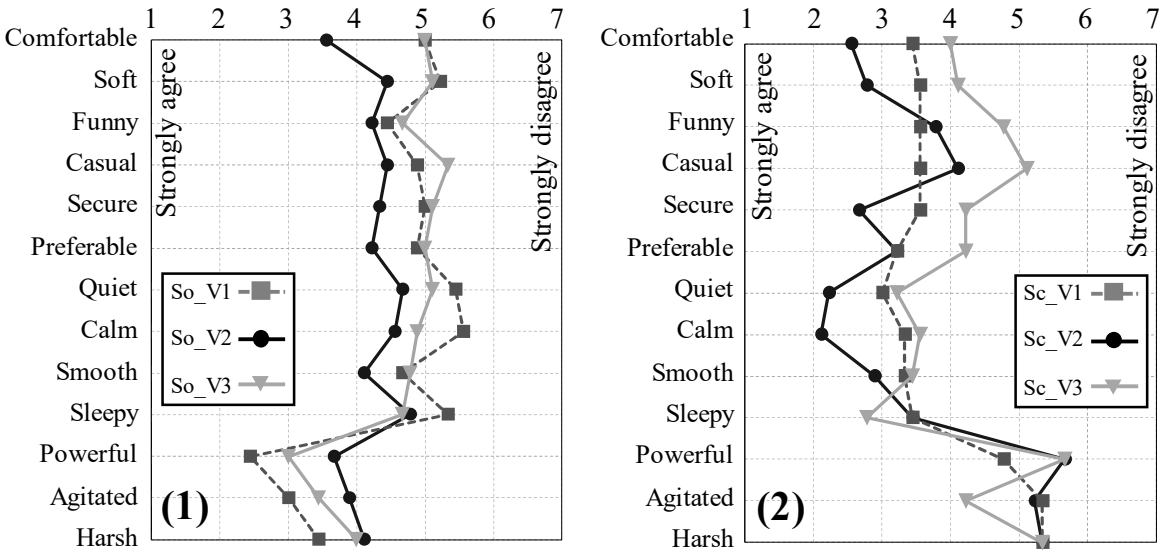

Figure 8 – Results of the SD test.

4. CONCLISION

The effect of additionally presented visual stimuli on the subjective impression of the sound environment inside the rooms located along vehicle road was investigated by combining the auralization scheme of the transmitted pass-by sound of running vehicle and the visualization scheme of inside the room. It has been indicated that the evaluated results of the sound impression for the pass-by vehicle sound transmitted into the room varied depending on the presented visual stimuli.

REFERENCES

[1] Kuwano S, Namba S, Hayashi Y, Komatsu M, Kato T, Auditory and visual interaction in the aesthetic evaluation of environment, Emp. Stud. Arts 2001; 19(2): 191–200.

[2] Preis A, Kociński J, Hafke-Dys H, Wrzosek W, Audio-visual interactions in environment assessment, Sci. Total Environ. 2015; 523:191–200.

[3] Viollon S, Lavandier C, Drake C, Influence of visual setting on sound ratings in an urban environment, Appl. Acoust. 2002; 63(5): 493–511.

[4] Szychowska M, Hafke-Dys H, Preis A, Kociński J, Kleka P, The influence of audio-visual interactions on the annoyance ratings for wind turbines, Appl. Acoust. 2017; 129(1): 190–203.

[5] Bangjun Z, Lili S, Guoqing D, The influence of the visibility of the source on the subjective annoyance due to its noise, Appl. Acoust. 2003; 64(12):1205–15.

[6] Pedersen E, Larsman P, The impact of visual factors on noise annoyance among people living in the vicinity of wind turbines, J. Environ. Psychol. 2008; 28(4):379–89.

[7] Fujimoto K, Sakata N, Yada M, A comparison of annoyance of road traffic noise evaluated in the field with that in the laboratory, J. Arch. Plan. 1986; 364: 10–19.

[8] Tokunaga Y, Terashima T, Ishikawa A, Influence of visual information on subjective evaluation of road traffic noise, Inter-noise 2014; 3607–11.

[9] Fastl H, Audio-visual interactions in loudness evaluation, ICA 2004; 1–6.

[10] Kleiner M, Dalenbäck B, Svensson P. Auralization – an overview, J. Audio. Eng. Soc. 1993; 41(11): 861–75.

[11] Vorländer M. Building acoustics: from prediction models to auralization, Proc. acoustics 2006; 2006: 15–22.

[12] Vorländer M. Auralization, Berlin: Springer-Verlag; 2007.

[13] Thaden R. An algorithm for auralization of airborne sound insulation, J. Acoust. Soc. Am. 1999;105(2): 1259.
[14] Rindel JH, Holger J. Auralization of airborne sound insulation including the influence of source room, Proc. Euronoise 2006; 2006: 27.
[15] Tanaka M, Fujikawa T, Oshino Y, Tachibana H. Visual/aural simulation of road traffic noise, Proc. 18th ICA; 2004: 3023–4.
[16] Yokota T, Sakamoto S, Tachibana H. Visualization of sound propagation and scattering in rooms, Acoust. Sci. Tech. 2002; 23(1): 40–6.
[17] Asakura T, Miyajima T, Sakamoto S, Prediction method for sound from passing vehicle transmitted through building façade, Appl. Acoust., 2013; 74(5): 758–769.
[18] Asakura T, Tsujimura S, Yonemura M, Hyojin L, and Sakamoto S, Effect of immersive visual stimuli on the subjective evaluation of the loudness and annoyance of sound environments in urban cities, Appl. Acoust., 2019; 143: 141–150.